Copyright © 1981 American Telephone and Telegraph Company
THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 60, No. 6 July–August 1981
Printed in U.S.A.

Analysis of Maximum-Likelihood Sequence Estimation Performance for Quadrature Amplitude Modulation

By A. S. ACAMPORA

(Manuscript received December 10, 1980)

This paper considers maximum-likelihood sequence estimation (MLSE) for quadrature amplitude modulation (QAM) signaling at rates approaching several baud/Hz. In this regime, intersymbol interference and possibly cross channel coupling are the dominant transmission impairments. We derive the structure of a detector that optimally accommodates both impairments. A bit error rate performance bound is found, and the concept of an error state transition matrix is introduced to facilitate the analysis. We explore a modulation scheme wherein cross-channel coupling is intentionally introduced, and find that it improves detection efficiency. The use of MLSE may be an important consideration for power and spectrally efficient digital radio systems, either terrestrial or satellite, since rates approaching the Shannon limit may be attainable without channel coding, and frequency selective fading is handled in an optimum manner.

I. INTRODUCTION

The search for digital radio modulations which combine power and bandwidth efficiency with ease of implementation has attracted interest for many years.[1-9] Combined with the performance analysis of receivers which represent a compromise between optimality and practicality, this field remains the focus of much current research activity.[10-15] Quadrature amplitude modulation (QAM) is a particularly simple modulation to implement, and determination of the theoretical bit-error rate (BER) achievable for this modulation at an arbitrary rate of information transfer per unit bandwidth is the subject of this paper. Phase shift keying (PSK) is, of course, a special case of QAM.

The primary impairment to QAM transmission is intersymbol interference caused by bandlimiting at the transmitter and dispersion in

the channel. We investigate maximum-likelihood sequence estimation (MLSE) to optimally detect the impaired signal, which is also corrupted by additive Gaussian noise. In particular, we analyze the Viterbi algorithm for realizing MLSE and then proceed to find an upper bound for the BER. In this regard, the analysis to be presented simplifies, unifies, and expands earlier treatments of the same subject.[16,17] The primary goal is to obtain a basis against which the performance of any suboptimum receiver may be compared. In so doing, we show QAM to be capable of providing performance close to the Shannon limit, an observation of considerable importance for communication satellite and terrestrial digital radio applications. Additionally, study of the mechanisms responsible for MLSE error generation provides insight into appropriate waveshaping to improve BER performance. One such waveshaping technique, wherein the spectrum is intentionally asymmetric with respect to the transmitting filter passband (e.g., QAM-Single Sideband), is shown to generally improve the BER performance.

In Section II, motivation behind a study of QAM is illustrated, the QAM model is presented, and the MLSE algorithm is derived. Section III is devoted to the derivation of a BER outer bound, and in Section IV, this bound is applied to examples which demonstrate the inherent power of MLSE. We show that a BER of 10^{-4} can be maintained at a transmission rate of 5 bits/s/Hz with an energy per bit penalty no greater than 1 dB compared against ideal nonoverlapping rectangular signaling; a four-pole Butterworth transmit filter is assumed in this calculation.

II. QAM MODEL AND THE MLSE ALGORITHM

Figure 1a gives a general model of a digital communication system. Here, the source produces a stream of binary data $\mathbf{d} = \{d_0, d_1, \cdots, d_L\}$ which is converted into an analog waveform by the modulator for transmission to the destination. The modulator may include a band-limiting filter to contain spectral emission. White Gaussian noise is added in the channel, and it is the function of the receiver to reproduce a copy $\hat{\mathbf{d}}$ of the original data sequence with a reasonably low probability of bit error.

The output of the modulator can be represented by the expression

$$S_G(t) = \text{Re}\ \{[A(t, \mathbf{d}) + jB(t, \mathbf{d})]e^{-j\omega_0 t}\}, \tag{1}$$

where ω_0 is the carrier radian frequency and $A(t, \mathbf{d})$ and $B(t, \mathbf{d})$ are data-dependent real waveforms.

The form of $S_G(t)$ provides great flexibility to achieve both high power efficiency and high spectral efficiency. However, the modulator to generate the best $S_G(t)$ and the receiver to detect the data in the presence of corrupting noise may be quite difficult to implement, e.g.,

separate waveform synthesizers and receiver matched filters may be required for each of the 2^L possible outcomes of the data stream **d**.

To simplify the modulation and detection processes, we consider QAM, a special subset of (1) having the form

$$S(t) = \mathrm{Re}\left\{\sum_{k=0}^{N} [a_k(\mathbf{d}) + jb_k(\mathbf{d})]\bar{h}(t - kT)e^{-j\omega_0 t}\right\}, \tag{2}$$

where

$\bar{h}(t) = h_R(t) + jh_I(t)$ is a complex waveshape,

a_k and b_k are real numbers dependent upon the data,

T^{-1} is the channel signaling rate,

and

$N + 1$ is the number of complex channel symbols used to transmit the data **d**.

Equation (2) can readily be modified to admit the familiar staggered types of modulation for which the in-phase and quadrature pulses are offset from each other by the amount $T/2$.[6,9] Although MLSE can be applied to this case, we restrict our attention in the following to nonstaggered modulation. We expect that performance similar to what we report here is achievable with staggered modulation as well.

The waveform $S(t)$ may be generated as shown in Fig. 1b. Here, the numbers a_k and b_k are generated by a channel encoder and the

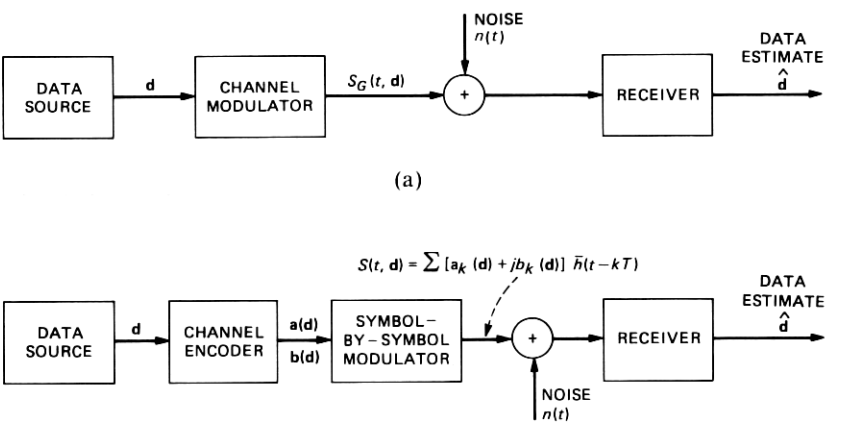

Fig. 1—Models for digital communication channels. (a) General digital channel. (b) Symbol-by-symbol QAM digital channel.

numbers a_k, b_k, a_{k+j}, b_{k+j} can be dependent. In the subsequent development we assume that, for $j \neq 0$, the pairs (a_k, b_k) and (a_{k+j}, b_{k+j}) are independent, corresponding to independent symbol-by-symbol signaling. The MLSE algorithm to be derived can be generalized to optimally detect $\mathbf{d}$ when the numbers a_k and b_k are dependent only upon the most recent K data symbols supplied to the encoder. A convolutional encoder is one such case which has been treated for real $h(t)$ and binary signaling.[18]

We note that the numbers a_k and b_k may be dependent. Thus, our treatment admits the familiar PSK type of modulation. Further, with no loss in generality, we assume in the following that $\bar{h}(t)$ is the combined impulse response of the transmitter and a possibly dispersive channel.

At the receiver, we observe

$$R(t) = S(t) + \text{Re}\,\{[n_1(t) + jn_2(t)]e^{-j\omega_0 t}\}, \tag{3}$$

where $n_1(t)$ and $n_2(t)$ are independent white Gaussian noise processes, each having two-sided spectral density $N_0/2$. Thus, the in-phase and quadrature components of $R(t)$ are as follows:

$$\begin{aligned}\mathcal{I}(t) &= \sum_{k=0}^{N} [a_k h_R(t - kT) - b_k h_I(t - kT)] + n_1(t) \\ &\triangleq S_R(t) + n_1(t)\end{aligned} \tag{4}$$

and

$$\begin{aligned}\mathcal{Q}(t) &= \sum_{k=0}^{N} [a_k h_I(t - kT) + b_k h_R(t - kT)] + n_2(t) \\ &\triangleq S_I(t) + n_2(t).\end{aligned} \tag{5}$$

Let

$$\mathbf{a} = \{a_0, a_1, \cdots, a_N\}, \qquad \mathbf{b} = \{b_0, b_1, \cdots, b_N\}. \tag{6}$$

Then, since the noise is white and Gaussian, the optimum receiver selects as detected data that sequence pair $(\hat{a}, \hat{b})$ for which the likelihood function,

$$L(\mathbf{a}, \mathbf{b}) = \exp\left(-\frac{2}{N_0}\int ([\mathcal{I}(t) - S_R(t)]^2 + [\mathcal{Q}(t) - S_I(t)]^2)\,dt\right), \tag{7}$$

is maximized. This is equivalent to maximizing, with respect to the hypothesis $(\mathbf{a}, \mathbf{b})$, the quantity

$$\begin{aligned}\Lambda(\mathbf{a}, \mathbf{b}) = 2\int_{-\infty}^{\infty} &[\mathcal{I}(t)S_R(t; \mathbf{a}, \mathbf{b}) + \mathcal{Q}(t)S_I(t; \mathbf{a}, \mathbf{b})]\,dt \\ &- \int_{-\infty}^{\infty} [S_R^2(t; \mathbf{a}, \mathbf{b}) + S_I^2(t; \mathbf{a}, \mathbf{b})]\,dt.\end{aligned} \tag{8}$$

From (4) and (5),

$$\Lambda(\mathbf{a}, \mathbf{b}) = 2 \sum_{k=0}^{N} \int_{-\infty}^{\infty} \mathscr{I}(t)[a_k h_R(t - kT) - b_k h_I(t - kT)]dt$$

$$+ 2 \sum_{k=0}^{N} \int_{-\infty}^{\infty} \mathscr{Q}(t)[a_k h_I(t - kT) + b_k h_R(t - kT)]dt$$

$$- \sum_{k=0}^{N} \sum_{m=0}^{N} \int_{-\infty}^{\infty} [a_k h_R(t - kT) - b_k h_I(t - kT)]$$

$$\cdot [a_m h_R(t - mT) - b_m h_I(t - mT)]dt$$

$$- \sum_{k=0}^{N} \sum_{m=0}^{N} \int_{-\infty}^{\infty} [a_k h_I(t - kT) + b_k h_R(t - kT)]$$

$$\cdot [a_m h_I(t - mT) + b_m h_R(t - mT)]dt. \tag{9}$$

Let

$$\alpha_k = \int_{-\infty}^{\infty} \mathscr{I}(t) h_R(t - kT)\, dt, \qquad \beta_k = \int_{-\infty}^{\infty} \mathscr{I}(t) h_I(t - kT)\, dt,$$

$$\gamma_k = \int_{-\infty}^{\infty} \mathscr{Q}(t) h_R(t - kT)\, dt, \qquad \rho_k = \int_{-\infty}^{\infty} \mathscr{Q}(t) h_I(t - kT)\, dt. \tag{10}$$

Then, the quantities in (10) can be generated by processing both $\mathscr{I}(t)$ and $\mathscr{Q}(t)$ through filters matched to $h_R(t)$ and $h_I(t)$, and sampling the output at time $t = kT$.

Let

$$y_k = \alpha_k + \rho_k, \qquad z_k = \gamma_k - \beta_k. \tag{11}$$

Finally, let us define

$$\chi_{k-m} = \int_{-\infty}^{\infty} [h_R(t - kT)h_R(t - mT) + h_I(t - kT)h_I(t - mT)]dt, \tag{12}$$

$$\zeta_{k-m} = \int_{-\infty}^{\infty} [h_I(t - kT)h_R(t - mT) - h_R(t - kT)h_I(t - mT)]dt. \tag{13}$$

Then

$$\Lambda(\mathbf{a}, \mathbf{b}) = 2 \sum_{k=0}^{N} (a_k y_k + b_k z_k) - \sum_{k=0}^{N} \sum_{m=0}^{N} (a_k a_m + b_k b_m) \chi_{k-m} - \sum_{k=0}^{N} \sum_{m=0}^{N} (a_k b_m - b_k a_m) \zeta_{k-m}. \tag{14}$$

Thus, for each hypothesis under test, the maximum-likelihood receiver forms a linear combination of the received statistics from which is subtracted a stored constant which is independent of the received data. (This constant must be slowly updated if the channel frequency characteristics change as during frequency-selective fading. Our analysis applies to each instantaneous characteristic since the detection epoch is far smaller than the time intervals associated with such changing characteristics.)

We note that the matched filtering operation can be performed in the passband, and that the received statistics y_k and z_k are, respectively, the in-phase and quadrature samples of the passband matched filter response to $\bar{h}(t)$,

$$\bar{v}(t) = \frac{1}{2\pi} \int |H(\omega)|^2 e^{j\omega t} d\omega = \int \bar{h}(\tau) \bar{h}^* (\tau - t) d\tau. \tag{15}$$

Then

$$\bar{v}(nT) = \chi_n + j\zeta_n, \tag{16}$$

and we see that the quantities χ_n and ζ_n are the real and imaginary samples of the matched filter response to $h(t)$ at time nT. These are, then, real and imaginary components of the intersymbol interference. From (12) and (13),

$$\chi_{-n} = \chi_n, \qquad \zeta_{-n} = -\zeta_n, \qquad \zeta_0 = 0. \tag{17}$$

In the following, we assume that the intersymbol interference vanishes for $|n| > M$, a positive integer.

For QAM, we can use the Viterbi algorithm to maximize (14) with respect to $(\mathbf{a}, \mathbf{b})$. Rewriting (14),

$$\Lambda(\mathbf{a}, \mathbf{b}) = \sum_{k=0}^{N} a_k \left[2y_k - a_k \chi_0 - 2 \sum_{m=k-M}^{k-1} (a_m \chi_{k-m} + b_m \zeta_{k-m}) \right] + \sum_{k=0}^{N} b_k \left[2z_k - b_k \chi_0 - 2 \sum_{m=k-M}^{k-1} (b_m \chi_{k-m} - a_m \zeta_{k-m}) \right], \tag{18}$$

where the convention is adopted that a_k and b_k are identically zero for $k < 0$.

Let the partial sum $\Lambda_n(\mathbf{a}, \mathbf{b})$ be the first n terms in each of the sums appearing in eq. (18). Then

$$\Lambda_n(\mathbf{a}, \mathbf{b}) = \Lambda_{n-1}(\mathbf{a}, \mathbf{b}) + a_n \left[2y_n - a_n\chi_0 - 2 \sum_{m=n-M}^{n-1} (a_m\chi_{n-m} + b_m\zeta_{n-m}) \right] + b_n \left[2z_n - b_n\chi_0 - 2 \sum_{m=n-M}^{n-1} (b_m\chi_{n-m} - a_m\zeta_{n-n}) \right]. \tag{19}$$

Equation (19) is in a form such that the Viterbi algorithm can be directly applied. We see that to calculate $\Lambda_n(\mathbf{a}, \mathbf{b})$, we must add to $\Lambda_{n-1}(\mathbf{a}, \mathbf{b})$ a term which depends on the received data y_n and z_n and the hypothesis subvectors $(a_{n-M}, \cdots, a_n)$ and $(b_{n-M}, \cdots, b_n)$. Furthermore, the next calculation to find $\Lambda_{n+1}(\mathbf{a}, \mathbf{b})$ no longer depends on a_{n-M} and b_{n-M}. Thus, we can perform MLSE by means of a trellis diagram containing I^M states, where I is the number of discrete values which can be assumed by each symbol pair (a_n, b_n). Let us label each state by the $2M$-tuple $a_{j-M+1}, \cdots, a_j, b_{j-M+1}, \cdots, b_j$. Then, permissible state transitions take the form shown in Fig. 2, drawn for (a_n, b_n) independent and binary (hence, $I = 4$) and $M = 2$. For each permissible transition, we can compute the branch metric $\Lambda_j(\mathbf{a}, \mathbf{b}) - \Lambda_{j-1}(\mathbf{a}, \mathbf{b})$ [see eq. (19)].

Maximum-likelihood sequence estimation is performed via recursive application of (19) for $n = 0, 1, \cdots, j, \cdots, N$. Suppose the partial sum of the most likely path leading into each state for $n = j - 1$ is known. We calculate the partial sum of the I competing paths leading into each state for $n = j$ by adding the partial sum of the most likely path into each state at $n = j - 1$ to the branch metric corresponding to that transition. The largest of the partial sums among the paths merging at each state then corresponds to the most likely path leading to that state and is stored for future calculations. The hypothesis (a_{j-M}, b_{j-M}) corresponding to the surviving path at each state is also stored as part of the most likely path leading to that state since a_{j-M}, b_{j-M} does not affect subsequent calculations. Thus, at each node, $I - 1$ out of I possible paths are deleted from further consideration. The path through the trellis with the greatest metric Λ_N then identifies the most likely transmitted sequence $(\hat{\mathbf{a}}, \hat{\mathbf{b}})$.

Unlike the case of the general digital modulation of (1), the complexity of the *ML* receiver for QAM grows exponentially with M (the one-sided extent of the ISI) rather than exponentially with the message length.

III. BIT-ERROR RATE PERFORMANCE BOUND

To find an upper bound on the BER performance of MLSE, we invoke the concept of an error event. Consider an incorrect path which diverges from the correct path at depth p in the trellis and remerges with it for the first time at depth q. Since the correct and incorrect paths agree over the first p epochs, the difference in path metrics at the remerge point is given by

$$\Lambda_q - \Lambda'_q = \sum_{n=p+1}^{q} (a_n - a'_n) \Bigg[2y_n - (a_n - a'_n)\chi_0 - 2 \sum_{m=n-M}^{n-1} \{(a_m - a'_m)\chi_{n-m} + (b_m - b'_m)\zeta_{n-m}\} \Bigg] + \sum_{n=p+1}^{q} (b_n - b'_n) \Bigg[2z_n - (b_n - b'_n)\chi_0 - 2 \sum_{m=n-M}^{n-1} \{(b_m - b'_m)\chi_{n-m} - (a_m - a'_m)\zeta_{n-m}\} \Bigg], \quad (20)$$

where

Λ_q is the metric of the correct path,

Λ'_q is the metric of the remerging incorrect path,

a_n, b_n are the channel symbols along the correct path,

and

a'_n, b'_n, are the channel symbols along the incorrect path.

Let

$$E_n = \tfrac{1}{2}(a_n - a'_n), \quad (21)$$

$$F_n = \tfrac{1}{2}(b_n - b'_n). \quad (22)$$

Then, substituting (10) and (11) into (20), and after extensive simplification (see Ref. 18 for a similar derivation), we obtain the result

$$\Lambda_q - \Lambda'_q = D + n_{eq}, \quad (23)$$

where

$$D = \sum_{n=p+1}^{q} E_n \Bigg[E_n + 2 \sum_{m=n-M}^{n-1} (E_m\chi_{n-m} - F_m\zeta_{n-m}) \Bigg] + \sum_{n=p+1}^{q} F_n \Bigg[F_n + 2 \sum_{m=n-M}^{n-1} (F_m\chi_{n-m} + E_m\zeta_{n-m}) \Bigg] \quad (24)$$

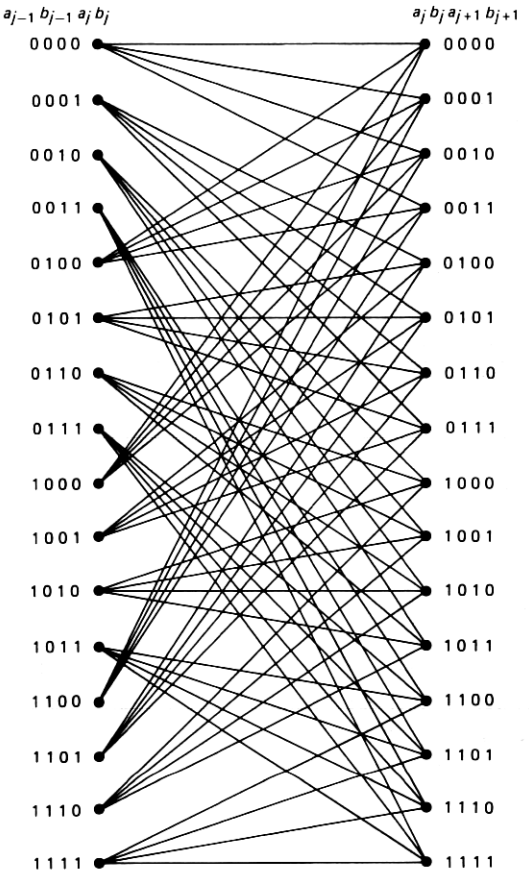

Fig. 2—State transition diagram for four-level QAM with four symbols of ISI.

and n_{eq} is a zero-mean Gaussian noise random variable with variance

$$\sigma^2 = DN_0/2. \tag{25}$$

In (24), we have assumed that the energy contained in $h(t)$ is unity.

At the qth epoch, the MLSE receiver must choose a survivor from among that path agreeing with the correct path over the most recent M epochs and all other incorrect paths remerging with it. An error will be committed if an incorrect path is chosen. Since the correct path may have been deleted by a prior decision, the metric comparisons may be between the remerging incorrect path and a partially correct path with metric greater than that of the overall correct path. The probability of committing an error at the qth epoch by selecting a particular one of the remerging incorrect paths of length $q - p$ is, then, overbounded for a particular correct sequence by

$$P_p \leq P\{\Lambda_q - \Lambda'_q < 0\}. \tag{26}$$

Recognizing that the true energy contained in $h(t)$ is $\chi_0 \triangleq e_b$, we now normalize the intersymbol interference coefficients by dividing each by χ_0 to obtain the result

$$P_p \leq Q(\sqrt{2e_bD/N_0}), \tag{27}$$

where $Q(\cdot)$ is the complementary error function.

Let N_p be the number of incorrect paths which diverge from the correct path at the pth epoch and remerge, for the first time, at the qth epoch, and let $\mathscr{E}_{p,j}$ be an error event associated with one such path. Then the probability of committing an error at the qth epoch is overbounded, for a particular correct path, by

$$P_E \leq P\left\{\bigcup_{p=1}^{q}\bigcup_{j=1}^{N_p}\mathscr{E}_{p,j}\right\}. \tag{28}$$

Applying the union bound,

$$P_E \leq \sum_{p=1}^{q}\sum_{j=1}^{N_p} P\{\mathscr{E}_{p,j}\}, \tag{29}$$

where, for a particular event $\mathscr{E}_{p,j}$, $P\{\mathscr{E}_{p,j}\}$ is the probability that $\Lambda_q - \Lambda'_q < 0$.

Let $\mu_{p,j}$ be the number of bit errors associated with the event $\mathscr{E}_{p,j}$. Then, the expected BER for decisions made at the qth epoch, conditioned on a particular correct sequence, is overbounded by:

$$P_B \leq \sum_{p=1}^{q}\sum_{j=1}^{N_p}\frac{\partial}{\partial\kappa}[\kappa^{\mu_{p,j}}P\{\mathscr{E}_{p,j}\}]\Bigg|_{\kappa=1}. \tag{30}$$

The particular form chosen for (30) will be useful later.

Finally, we note that the correct path is one of I^{q-p} outcomes over the unmerged span between epochs p and q. The result (30) must, therefore, be averaged over these outcomes to remove the conditional dependence on the correct path.

Thus, evaluation of the upper bound involves a weighted summation of the probabilities of all error events merging with the correct path at the qth node. We will now develop a technique for performing this summation in an orderly fashion.

We note from (25) that for every error event, $D > 0$. Also, for each error event, D is dependent only upon the error sequences (E_n, F_n), $p < n \leq q$. For one particular error event, the quantity D will assume its minimum value D_{min}. Then, using (27) and the inequality[1] $Q(\sqrt{x+y}) \leq (\sqrt{x})e^{-y/2}$ for $x > 0$, $y > 0$, we obtain

$$P_B \leq Q\left(\sqrt{\frac{2e_bD_{\min}}{N_0}}\right)\exp\{e_bD_{\min}/N_0\} \cdot \sum_{p=1}^{q}\sum_{j=1}^{N_p}\frac{\partial}{\partial\kappa}[\kappa^{\mu_{p,j}}\exp\{-e_bD_{p,j}/N_0\}]. \tag{31}$$

Hence, evaluation of the upper bound involves computation of the set of Euclidean distances $D_{p,j}$ for each correct path. The error-state transition matrix, to be described, provides a systematic way to find these distances, average over correct paths, and perform the sum.

Referring to eq. (24), we see that the contribution to the distance between a correct and an incorrect path incurred between nodes j and $j+1$, where $p \le j \le q$, is given by

$$d = \sqrt{2e_b/N_0}\, E_j \left[E_j + 2 \sum_{m=j-M}^{j-1} (E_m \chi_{j-m} + F_m \zeta_{j-m}) \right] + \sqrt{2e_b/N_0}\, F_j \left[F_j + 2 \sum_{m=j-M}^{j-1} (F_m \chi_{j-m} - E_m \zeta_{j-m}) \right]. \quad (32)$$

This contribution can be represented as the distance incurred by a transition from error state $E_{j-M}, E_{j-M+1}, \cdots, E_{j-1}, F_{j-M}, F_{j-M+1}, \cdots, F_{j-1}$ to error state $E_{j-M+1}, E_{j-M+2}, E_j, F_{j-M+1}, F_{j-M+2}, \cdots, F_j$. A particular pair of elements (E_n, F_n) from an error sequence may assume no more than $J \le I(I-1) + 1$ discrete values. For example, if a_n and b_n are binary and independent, then the pair (E_n, F_n) may assume the nine values: (0, 0), (0, −1), (0, 1), (1, −1), (1, 0), (1, 1), (−1, −1), (−1, 0), (−1, 1). Excluding the all zeros error state, $J^M - 1$ error states remain. We write down a $(J^M - 1) \times (J^M - 1)$ error-state transition matrix $\mathbf{T} = \{t_{i,l}\}$, where the indices i and l each range over the $J^M - 1$ integers labeling the error states. The element $t_{i,l}$ represents the weighted distance incurred by the $i \to l$ error-state transition:

$$t_{i,l} = f_{i,l} \kappa^{\mu_{i,l}} e^{d_{i,l}},$$

and

$d_{i,l}$ is the Euclidean distance incurred in permissible transitions $i \to l$;

$\mu_{i,l}$ is the number of bit errors incurred in the $i \to l$ transition;

$f_{i,l}$ is a factor representing the fraction of data pairs (a_j, b_j) which can produce the error (E_j, F_j) at any epoch j.

Element $t_{i,l} = 0$ for forbidden transitions.

The weighted distance incurred along any error path between epochs j_1 and j_2 can be uniquely determined from either the error sequence between $j_1 - M$ and j_2 or the error-state transitions between j_1 and j_2. Using the latter approach, the weighted distance incurred along an error path between epochs j_1 and j_2 is simply the product of the weighted transition distances $t_{i,l}$ along that path.

As an example of the values to be assigned to the factors $f_{i,l}$ and $\mu_{i,l}$, suppose that the lth error state at epoch j is characterized by

$(E_j, F_j) = (1, -1)$, and the channel symbols a_j and b_j are binary and independent. Then, a_j and b_j must equal $+1$ and -1, respectively, implying that $f_{i,l} = \frac{1}{4}$, and the number of errors incurred $\mu_{i,l} = 2$. Alternatively, if $(E_j, F_j) = (1, 0)$, then a_j must equal $+1$, b_j can be ± 1, $f_{i,l} = \frac{1}{2}$, and $\mu_{i,l} = 1$.

Since each error state can be reached from no more than $J - 1$ error states, each row in the error-state transition matrix contains no more than $J - 1$ nonzero entries, and $\mathbf{T}$ is quite sparse.

The error states are numbered from 1 to $J^M - 1$. Then, at any epoch $j - 1$, an error event may begin by a transition from the correct state $(E_{j-M} = 0, \cdots, E_{j-1} = 0, F_{j-M} = 0, \cdots, F_{j-1} = 0)$ to one of the $J - 1$ states $(E_{j-M+1} = 0, \cdots, E_{j-1} = 0, E_j, F_{j-m+1} = 0, \cdots, F_{j-1} = 0, F_j)$. We introduce the $J^M - 1$ dimensional column vector $\mathbf{V}_q$. The nth element of $\mathbf{V}_q$ is the accumulated weighted distances of all error paths spanning q epochs which merge with the nth error state at any node. Then, $\mathbf{V}_1$ contains $J - 1$ nonzero elements, corresponding to the $J - 1$ error states given above. The values assumed by these $J - 1$ elements are

$$t_i = \sqrt{2e_b/N_0\,[E_j^2 + F_j^2]\kappa^{\mu_i} f_i}, \tag{33}$$

where

μ_i is the number of bit errors incurred in transitioning from the all-zeros error state to error state i;

f_i is the fraction of data pairs (a_j, b_j) capable of producing a transition from the all-zeros error state to error state i;

and

i can assume one of the $J - 1$ values corresponding to the indices of error states which can be reached from the all-zeros error state.

Also, the vector $\mathbf{V}_q$ is given by

$$\mathbf{V}_q = \mathbf{T}^{(q-1)}\mathbf{V}_1. \tag{34}$$

At depth p, a remerge may occur from the $J - 1$ error states $(E_j, E_{j+1} = 0, \cdots, E_{j+p-1} = 0, F_j, F_{j+1} = 0, \cdots, F_{j+p-1} = 0)$. We introduce a second $J^M - 1$ dimensional vector $\mathbf{U}$ containing only $J - 1$ nonzero elements. The nth element of $\mathbf{U}$ is equal to unity only if error state n merges with the all-zeros error state. Then the contribution to the BER bound at depth p from the point of divergence is given by

$$P_{B,p} \leq \mathbf{U}^T \cdot \left[\frac{\partial}{\partial \kappa}\,(\mathbf{T}^{p-1}\mathbf{V}_1)\right]\Bigg|_{\kappa=1}. \tag{35}$$

Thus, using the error-state transition matrix, the BER bound (30)

can be averaged over the correct sequence and rewritten as:

$$P_B \leq \mathbf{U}^T \cdot \sum_{p=1}^{\infty} \left[\frac{\partial}{\partial \kappa} (\mathbf{T}^{p-1}\mathbf{V}_1) \right] \Bigg|_{\kappa=1} \cdot Q\left(\sqrt{\frac{2e_b D_{\min}}{N_0}} \right) \exp(e_b D_{\min}/N_0). \tag{36}$$

An alternate form is

$$P_B \leq \mathbf{U}^T \cdot \left\{ \lim_{\epsilon \to 0} \frac{1}{\epsilon} \sum_{p=1}^{\infty} [\mathbf{T}^{p-1}(\kappa = 1 + \epsilon)\mathbf{V}_1 - \mathbf{T}^{p-1}(\kappa = 1)\mathbf{V}_1] \cdot Q\left(\sqrt{\frac{2e_b D_{\min}}{N_0}} \right) \exp(e_b D_{\min}/N_0) \right\}. \tag{37}$$

The bound will converge if the matrix series

$$\mathbf{S} = \sum_{p=1}^{\infty} \mathbf{T}^{p-1}(\kappa = 1) \tag{38}$$

converges. If (38) converges, then

$$\mathbf{S} = (\mathbf{I} - \mathbf{T})^{-1}, \tag{39}$$

and

$$P_B \leq \mathbf{U}^T \cdot \left[\frac{\partial}{\partial \kappa} (\mathbf{I} - \mathbf{T})^{-1} \mathbf{V}_1 \right] \Bigg|_{\kappa=1} \cdot Q(\sqrt{2e_b D_{\min}/N_0}) \exp(e_b D_{\min}/N_0). \tag{40}$$

A necessary and sufficient condition for the convergence of $\mathbf{S}$ is that the largest eigenvalue of $\mathbf{T}$ have magnitude less than unity.

If $J^M - 1$ is small, then the largest eigenvalue of $\mathbf{T}$ can actually be found, and if the convergence criterion is satisfied, then the BER bound for MLSE can be found in closed form by inverting $\mathbf{I} - \mathbf{T}$ and substituting into (40). When $J^M - 1$ is too large to conveniently admit such closed-form solution, then mathematical rigor must be sacrificed by returning to (37), and truncating the series at some point $p_{\max}$ such that the remainder may safely be assumed to be small. In this event, the order of matrix $\mathbf{T}$ is usually quite large, and some concern may arise from the apparent need to repetitively multiply large matrices. This is not the case, however, since the term $\mathbf{T}^{p-1}\mathbf{V}_1$ may be found recursively as

$$\mathbf{T}^{p-1}\mathbf{V}_1 = \mathbf{T}(\mathbf{T}^{p-2}\mathbf{V}_1), \tag{41}$$

i.e., the term $\mathbf{T}^{p-1}\mathbf{V}_1$ may be found by recursively multiplying the matrix $\mathbf{T}$ by a vector.

IV. APPLICATIONS

As a first application, we consider MLSE detection of 4ϕ PSK with overlapping rectangular pulses. Let

$$\bar{h}(t) = p(t)e^{j\Delta t}, \tag{42}$$

where

$$p(t) = \begin{cases} \sqrt{2e_b/\tau}, & 0 \le t < \tau \\ 0, & \text{elsewhere.} \end{cases} \tag{43}$$

Then,

$$S(t) = \text{Re}\left\{ \sum_{k=0}^{L} (a_k + jb_k)\, \bar{h}(t - kT)e^{-j\omega_0 t} \right\}, \tag{44}$$

where $T < \tau$ is the signaling interval, and a_k and b_k are independent and equally likely to be ± 1. Suppose $T = \tau/4$, i.e., 8 bits are transmitted every symbol duration (τ seconds). If $\Delta = 0$, the four possible signal points for each transmitted symbol are as shown in Fig. 3a, corresponding to the familiar 4ϕ PSK modulation. Similarily, if $\Delta = 2\pi/\tau$, then $\bar{h}(t - kT) = p(t - kT)\exp[j\Delta(t - kT) = p(t - kT)\exp(-j^{\pi k/2}) \cdot \exp(j^{2\pi t/\tau})$, and the four signal points for each symbol transmitted are again as shown in Fig. 3a, although the points exhibit a progressive 90-degree phase rotation for each successive symbol transmitted. The carrier is also offset to the new value $\omega' = \omega_0 + 2\pi/\tau$. However, since a_k and b_k are independent and equally likely to be ± 1, the error events for the waveshapes corresponding to $\Delta = 0$ and $\Delta = 2\pi/\tau$ are identical, and both must, therefore, exhibit the same detection performance.

For $\Delta = \pi/\tau$, $\bar{h}(t - kT) = p(t - kT)e^{-j^{\pi k/4}}e^{-j^{\pi t/\tau}}$, the carrier is offset to the new value $\omega' = \omega_0 + \pi/\tau$, and the signal points alternate on

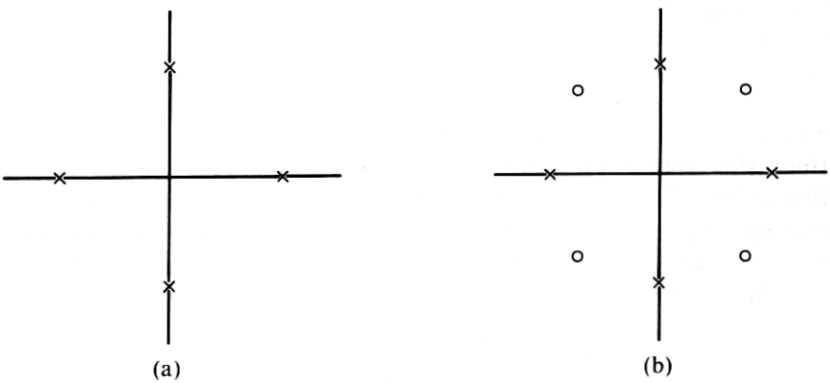

Fig. 3—Signal constellations for four-level QAM. (a) 4ϕ PSK; carrier offset $\Delta = 0$ or π/τ. (b) 4ϕ PSK; carrier offset $\Delta = \pi/\tau$. Signal constellation alternates between crosses and circles.

successive symbol transmissions between the two sets of locations shown in Fig. 3b. As will subsequently be shown, detectability for $\Delta = \pi/\tau$ is superior to that for $\Delta = 0$.

For rectangular signaling with $T = \tau/4$, it is readily established that $M = 3$, and the MLSE receiver for 4ϕ PSK contains $4^3 = 64$ states. The error-state transition matrix needed for analysis, however, is of the order $N = 9^3 - 1 = 728$. We number the error states by means of the six-tuple $(E_1, F_1, E_2, F_2, E_3, F_3)$, where each element can assume 3 values $(0, \pm 1)$. The state where each element is equal to zero corresponds to a merge and is excluded from the error-state transition matrix. For numerical purposes, a convenient way to number the error states is given by

$$B = \begin{cases} 1 + 3^0(F_3 + 1) + 3^1(E_3 + 1) + 3^2(F_2 + 1) \\ \quad +3^3(E_2 + 1) + 3^4(F_1 + 1) + 3^5(E_1 + 1), \\ \qquad\qquad\qquad\qquad B \le (9^M - 1)/2, \\ 3^0(F_3 + 1) + 3^1(E_3 + 1) + 3^2(F_2 + 1) \\ \quad +3^3(E_2 + 1) + 3^4(F_1 + 1) + 3^5(E_1 + 1), \\ \qquad\qquad\qquad\qquad B > (9^M - 1)/2, \end{cases} \tag{45}$$

where B is the state number. If the above numbering scheme for a_k, b_k binary and independent is generalized to arbitrary M, then it can readily be established that the error-state transition matrix possesses the symmetry:

$$t_{i,j} = t_{P+1-i,P+1-j}, \qquad P = 9^M - 1. \tag{46}$$

Each row in $\mathbf{T}$ contains no more than nine nonzero elements, corresponding to the permissible transitions.

For the case $\Delta = 0$, $D_{\min}$ was found by search to be 0.5, and for $\Delta = \pi/\tau$, $D_{\min}$ was found to be 0.757. Bit-error rate upper bounds for the two cases as functions of e_b/N_0 are shown in Fig. 4. Also shown, for comparison, is the ideal BER for quadraphase signaling with nonoverlapping rectangular pulses, i.e., 2 bits transmitted per symbol duration T. We see that, for $10^{-6} \le P_B \le 10^{-4}$, MLSE performance for $t = \tau/4$, $\Delta = \pi/\tau$ is within 1.5 to 2 dB of that achieved with nonoverlapping pulses, although the signaling rate is four times as high. Also, over the same BER range, performance of signaling with $\Delta = 0$ is 1.5 to 2 dB poorer than with $\Delta = \pi/\tau$.

We next consider the transmission of overlapping pulses of finite duration with amplitude tapering to contain spectral emission. The pulse $h(t)$ takes the form:

$$\bar{h}(t) = \begin{cases} [1 + \epsilon \cos(2\pi t)/\tau]\exp(j^{\Delta t}), & -\tau/2 < t \le \tau/2, \\ 0, \quad \text{elsewhere}, \end{cases} \tag{47}$$

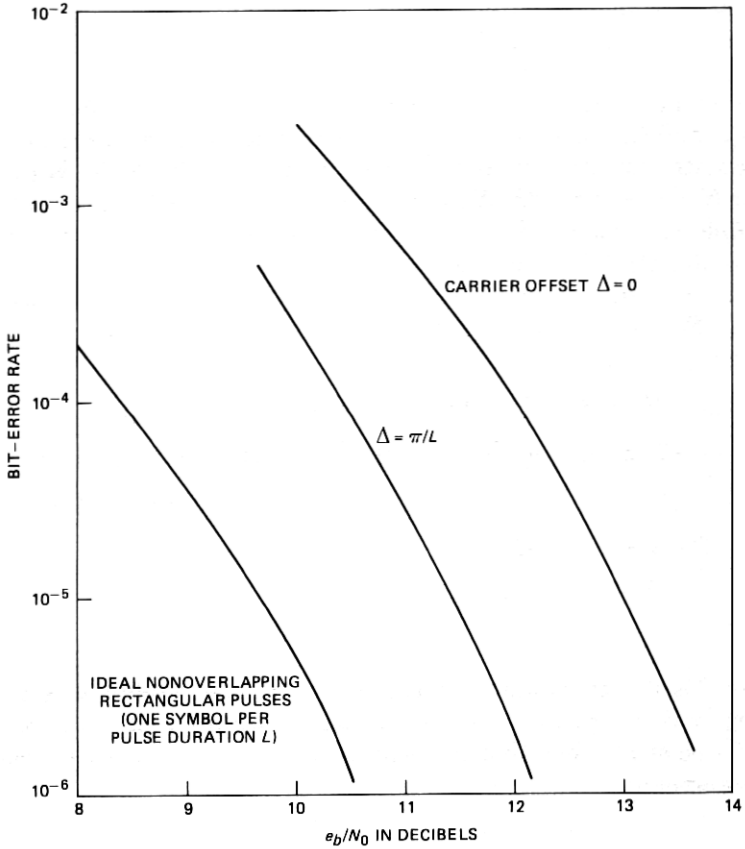

Fig. 4—Bit-error rate vs e_b/N_0 for overlapping rectangular pulses, with no carrier offset and with optimum carrier offset. Four symbols are transmitted per pulse duration L.

where $0 < \epsilon < 1$. The transmitted signal is given by (44) with $\bar{h}(t)$ as above, and the power spectrum for various values of ϵ, with $\Delta = 0$, are as shown in Fig. 5. As ϵ increases in amplitude, the main lobe of the power spectrum tends to increase, but the sidelobes are reduced.

For the particular case $\epsilon = 0.45$, the peak sidelobe is down 24 dB relative to the main lobe height. Let us define the R.F. bandwidth as being twice the baseband frequency at which the power spectrum dips permanently below −24 dB. Then, from Fig. 5, the spectral efficiency of the transmission (bit rate/R.F. bandwidth) is equal to $(2/2.5)(\tau/T)$ bits/s/Hz. where, as before, T^{-1} is symbol signaling rate.

The MLSE error rate bounds for this pulse appear in Fig. 6. For $T = \tau/4$, corresponding to $M = 3$, it was found that $D_{\min}$ occurs at $\Delta = \pi/2\tau$, and for $T = \tau/5$, corresponding to $M = 4$, it was found that $D_{\min}$ occurs

at $\Delta = 5\pi/8\tau$. Both cases are compared against $\Delta = 0$. Results for $T = \tau/4$, corresponding to a transmission rate of 3.2 bits/s/Hz, show that for $P_B = 10^{-4}$, optimum carrier offset is about 0.5 dB superior to no offset, and that e_b/N_0 must increase by about 4 dB, compared against nonoverlapping pulses, to maintain the error rate. However, for the same pulse shape, transmission of nonoverlapping pulses provides a rate of only 0.8 bits/s/Hz. Thus, the transmission rate is higher by a factor of four. For $T = \tau/5$, the transmission rate is 4 bits/s/Hz, (a factor of five higher than for nonoverlapping pulses), and the required increase in e_b/N_0 is about 5.1 dB for optimum carrier offset. Performance without carrier offset is, again, about 0.5 dB inferior.

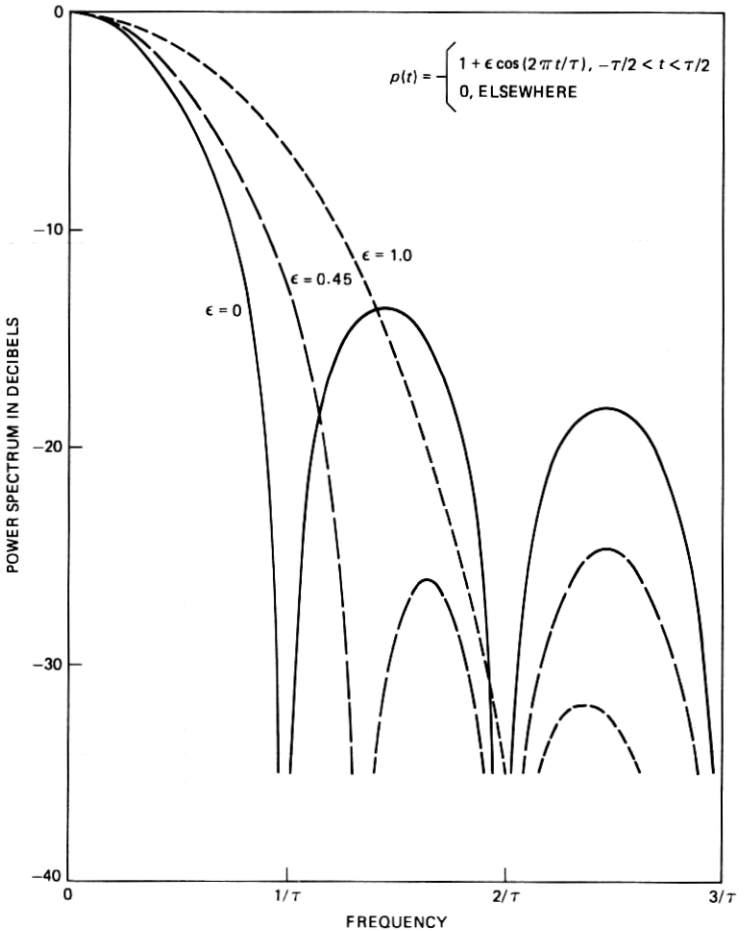

Fig. 5—Spectrum of raised-cosine pulses. $p(t) = 1 + \epsilon\cos(2\pi t/\tau)$ for $-\tau/2 < t < \tau/2$, and $p(t) = 0$ elsewhere.

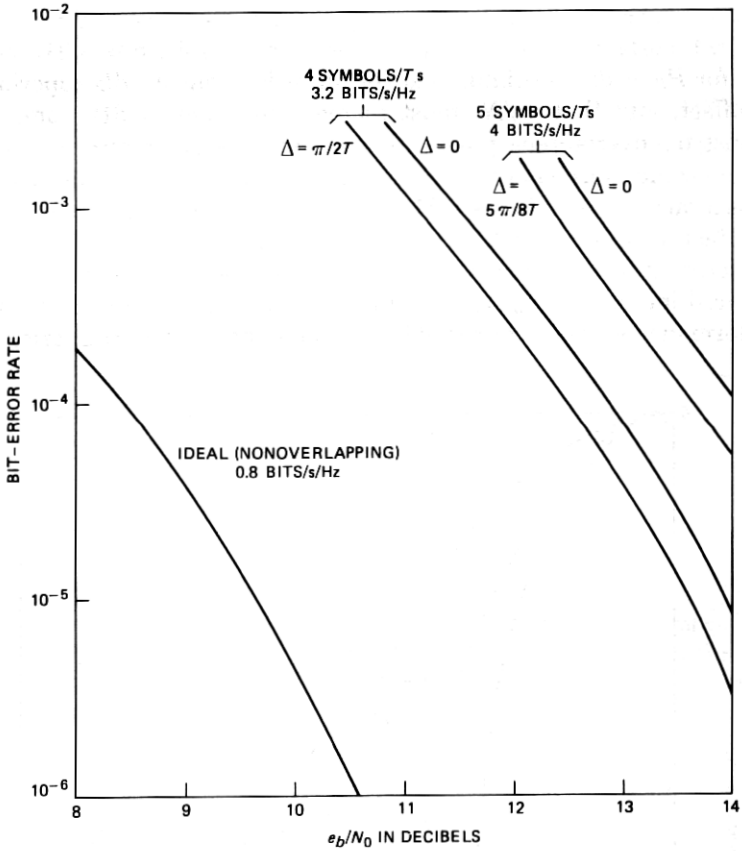

Fig. 6—Bit-error rate for overlapping raised-cosine pulses with $\epsilon = 0.45$, for no carrier offset and optimum carrier offset.

Let us compare these results against M-ary PSK signaling with nonoverlapping pulses. We define the carrier-to-noise (CNR) ρ required to achieve a given BER by

$$\rho = Re_b/N_0, \tag{48}$$

where R is the transmission rate in bits/s/Hz. For nonoverlapping binary pulses (4ϕ-PSK), the transmission rate is 0.8 bits/s/Hz, and to maintain a BER of 10^{-4}, $e_b/N_0 = 8.4$ dB and $\rho = 7.4$ dB. If we double the transmission rate to 1.6 bits/s/Hz by employing 16ϕ PSK with nonoverlapping pulses, the required CNR ratio increases by 12 dB to $\rho = 19.4$ dB. By contrast, MLSE with optimum carrier offset provides a transmission rate four times higher (3.2 bits/s/Hz) with $\rho = 17$ dB, and a transmission rate five times higher (4 bits/s/Hz) with $\rho = 19.7$ dB. Thus, for $\rho = 19.7$ dB, quadraphase modulation with overlapping

pulses, optimum carrier offset, and MLSE can more than double the capacity compared against 16ϕ PSK with nonoverlapping pulses.

As a final example, we give perhaps our most dramatic demonstration of the inherent power of MLSE. We consider that $\bar{h}(t)$ is the impulse response of a four-pole Butterworth filter whose center frequency is offset from the carrier frequency. At baseband, the matched-filtered response to $\bar{h}(t)$ has a spectrum

$$H(\omega) = \left[1 + \left(\frac{\omega - \omega_1}{\omega_{3\text{dB}}}\right)^8\right]^{-1}, \tag{49}$$

where $\omega_1 = \alpha\omega_{3\text{dB}}$ is the carrier offset, and $\omega_{3\text{dB}}$ is the cutoff frequency of the Butterworth filter. We arbitrarily define the R.F. bandwidth as twice the 3 dB cutoff frequency for no carrier offset. Transmission of $S(t)$, (44), at a rate $T^{-1} = \omega_{3\text{dB}}/2\pi$ then corresponds to $R = 1$ bit/s/Hz. We consider transmission of overlapping pulses at a rate five times as high. The various χ_n's and ζ_n's are found from (15) and (16) via contour integration. The impulse response is truncated at $M = 4$, and $|\chi_n|$ and $|\zeta_n|$ are both less than 0.02 for $n > 4$.

Results are as shown in Fig. 7. The optimum carrier offset was found to be $\alpha = 0.625$, yielding $D_{\min} = 0.9$; for $\alpha = 0$, $D_{\min} = 0.8$. We see that, for both cases, results are close to ideal rectangular signaling with nonoverlapping pulses, and that the optimum carrier offset is superior to no offset by about 0.5 dB. To maintain a BER of 10^{-4}, $e_b/N_0 =$ 9.5 dB. Since the transmission rate is 5 bits/s/Hz, the required CNR is then $\rho = 16.5$ dB. At this CNR, Shannon's channel capacity $C/W = \log_2(1 + \rho)$ is 5.5 bits/s/Hz. Thus, for $P_e = 10^{-4}$, MLSE performance with the simple Butterworth signaling chosen is quite close to the Shannon limit. Even without carrier offset, the required CNR is about 17 dB, for which the Shannon limit is 5.7 bits/s/Hz.

V. CONCLUSION

Maximum-likelihood-sequence estimation has been shown to be a powerful technique to permit high rates of data transmission with straightforward channel signaling. In this treatment, channel coding techniques were not used, and Quadrature Amplitude Modulation was employed. Implementation of MLSE by means of the Viterbi Algorithm requires a detector with I^M states, where I is the number of levels which may be assumed by each channel symbol and M is the extent, in symbol intervals, of intersymbol interference. Complex signaling waveforms, which are generated by means of carrier offset modulation, can be accommodated and may improve detection efficiency; MLSE equalizes both the in-phase and quadrature components of the intersymbol interference.

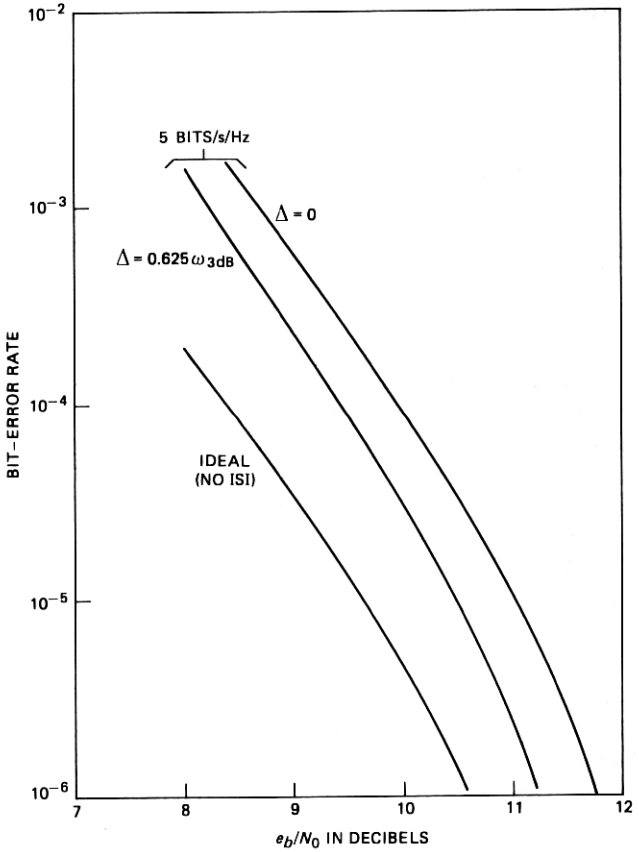

Fig. 7—Bit-error rate vs e_b/N_0 for four-pole Butterworth transmit filtering for no carrier offset and optimum carrier offset.

An error-state transition analysis was used to systematically compute a BER bound for MLSE. Applying this bound, it was shown by example that, with MLSE, transmission rates approaching the Shannon limit may be possible without channel coding.

The power of MLSE detection makes this technique a candidate for radio systems constrained in bandwidth. Since the MLSE approach admits complex waveshapes, channels admitting frequency-selective fading can readily be accommodated, provided that the slowly varying channel frequency characteristics are monitored to update the detection constants continuously. The real limitation on the application of MLSE involves the speed at which the detector can operate, and this places a constraint on the symbol rate of the radio channel. For rates under 30 Megabaud, MLSE may be quite feasible; implementation at real-time rates of 300 Megabaud, such as may apply to satellite

channels, is beyond the current state-of-the-art. However, if the satellite system employs time-division multiple access (TDMA), then the low-duty cycle for any ground station may permit MLSE implementation in nonreal time between TDMA burst arrivals. For duty cycles lower than 10 percent, this would appear to be feasible today. At higher duty cycles, nonreal time detection with parallel processors is another possibility.

REFERENCES

1. A. J. Viterbi, "Convolutional Codes and their Performance in Communication Systems," IEEE Trans. Commun. Tech., *COM-19* (October 1971), pp. 751–771.
2. J. A. Heller and I. M. Jacobs, "Viterbi Decoding for Satellite and Space Communication," IEEE Trans. Commun. Tech., *COM-19* (October 1971), pp. 835–848.
3. L. C. Tillotson, C. L. Ruthroff, and V. K. Prabhu, "Efficient Use of the Radio Spectrum and Bandwidth Expansion," Proc. IEEE, *61* (April 1973), pp. 445–452.
4. I. M. Jacobs, "Practical Applications of Coding," IEEE Trans. Inform. Theory, *IT-20* (May 1974), pp. 305–310.
5. V. K. Prabhu, "Spectral Occupancy of Digital Angle-Modulation Signals," B.S.T.J., *5* (April 1976), pp. 429–454.
6. S. A. Gronemeyer and A. L. McBride, "MSK and Offset QPSK Modulation," IEEE Trans. Commun., *COM-24* (August 1976), pp. 809–820.
7. V. K. Prabhu, "PSK-Type Modulation with Overlapping Baseband Pulses," IEEE Trans. Commun., *COM-25* (September 1977), pp. 980–990.
8. F. deJager and C. B. Dekker, "Tamed Frequency Modulation, A Novel Method to Achieve Spectrum Economy in Digital Transmission," IEEE Trans. Commun., *COM-26* (May 1978), pp. 534–542.
9. W. J. Weber III, P. H. Stanton, and J. T. Sumida, "A Bandwidth Compressive Modulation System Using Multi-Amplitude Minimum Shift Keying (MAMSK)," IEEE Trans. Commun., *COM-26* (May 1978), pp. 543–551.
10. S. Benedetto, G. DeVincentiis, and A. Luvison, "Error Probability in the Presence of Intersymbol Interference and Additive Noise for Multilevel Digital Signals," IEEE Trans. Commun., *COM-21* (March 1973), pp. 181–190.
11. S. Benedetto and E. Biglieri, "On Linear Receivers for Digital Transmission Systems," IEEE Trans. Commun., *COM-22* (September 1974), pp. 1205–1215.
12. R. Fang and O. Shimbo, "Unified Analysis of a Class of Digital Systems in Additive Noise and Interference," IEEE Trans. Commun., *COM-22* (October 1973), pp. 1075–1091.
13. D. D. Taylor, H. C. Chan, and S. S. Haykin, "A Simulation Study of Digital Modulation Methods for Wideband Satellite Communications," IEEE Trans. Commun., *COM-24* (December 1976), pp. 1351–4.
14. K. Feher et al., "Comparison of PSK and Multilevel AM Data Modulation," Int. J. Electron., *41*, No. 2 (1976), pp. 153–8.
15. L. J. Greenstein and D. Vitello, "Digital Radio Receiver Responses for Combating Frequency—Selective Fading," IEEE Trans. Commun., *COM-27* (April 1979), pp. 671–681.
16. G. D. Forney, Jr., "Maximum-Likelihood Sequence Estimation of Digital Sequences in the Presence of Intersymbol Interference," IEEE Trans. Inform. Theory, *IT-18* (May 1972), pp. 363–378.
17. G. Ungerboeck, "Adaptive Maximum-Likelihood Receiver for Carrier Modulated Data Transmission Systems," IEEE Trans. Commun., *COM-22* (May 1974), pp. 624–636.
18. A. S. Acampora, "Maximum-Likelihood Decoding of Binary Convolutional Codes on Band-Limited Satellite Channels," IEEE Trans. Commun., *COM-26* (June 1978), pp. 766–776.